Published online February 11, 2004

Nucleic Acids Research, 2004, Vol. 32, No. 3 ***1075–1082***
DOI: 10.1093/nar/gkh263

Structure–function correlations derived from faster variants of a RNA ligase deoxyribozyme

Tracey K. Prior, Daniel R. Semlow, Amber Flynn-Charlebois, Imran Rashid and Scott K. Silverman*

Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA

Received November 17, 2003; Revised and Accepted January 13, 2004

ABSTRACT

We previously reported the *in vitro* selection of several Mg^{2+}-dependent deoxyribozymes (DNA enzymes) that synthesize a 2′–5′ RNA linkage from a 2′,3′-cyclic phosphate and a 5′-hydroxyl. Here we subjected the 9A2 deoxyribozyme to re-selection for improved ligation rate. We found two new DNA enzymes (7Z81 and 7Z48) that contain the catalytic core of 7Q10, a previously reported small deoxyribozyme that is unrelated in sequence to 9A2. A third new DNA enzyme (7Z101) is unrelated to either 7Q10 or 9A2. The new 7Z81 and 7Z48 DNA enzymes have ligation rates over an order of magnitude higher than that of 7Q10 itself and they have additional sequence elements that correlate with these faster rates. Our findings provide insight into structure–function relationships of catalytic nucleic acids.

INTRODUCTION

We have initiated *in vitro* selection efforts towards deoxyribozymes (DNA enzymes) that ligate RNA (1–4). Our previous studies have focused on two different ligation reactions: joining a 2′,3′-cyclic phosphate with a 5′-hydroxyl (1–3) and linking a 2′(3′)-hydroxyl with a 5′-triphosphate (4). For unclear mechanistic reasons, all of the currently known Mg^{2+}-dependent deoxyribozymes that mediate reaction of the 2′,3′-cyclic phosphate do so with formation of 2′–5′ phosphodiester linkages (**1**→**2** in Fig. 1), to the exclusion of isomeric 3′–5′ linkages that could also be formed (**1**→**3**) (1–3). Here we sought to optimize the RNA ligation rate (k_{obs}) of one of our originally reported 2′–5′ RNA ligase deoxyribozymes, 9A2 (1), by partial randomization and re-selection. Such efforts are important for practical purposes if deoxyribozymes are to be used as synthetic tools for RNA ligation. In addition, these experiments are significant for conceptual reasons to understand how DNA enzyme structure and function are interrelated.

From the selection experiments described here, we identified two new DNA enzymes (7Z81 and 7Z48) that have significantly increased k_{obs} relative to earlier RNA ligase deoxyribozymes such as 9A2. Sequence comparisons reveal that the new DNA enzymes are unrelated to 9A2 and instead have a catalytic core similar to the previously reported 7Q10, which was evolved from an RNA-cleaving DNA enzyme (2,3). The structural features of the new 7Q10-like deoxyribozymes correlate with functional differences in ligation activity, providing insight into structure–function relationships of catalytic nucleic acids.

MATERIALS AND METHODS

Nucleic acid reagents

Sources of materials are described in detail in our previous report (1). The left-hand RNA substrate sequence was 5′-UAAUACGACUCACUAUA-3′, with a 2′,3′-cyclic phosphate, and the right-hand RNA substrate sequence was 5′-GGAAGAGAUGGCGACGG-3′. For the ligation assays (see below), the left-hand RNA substrate was 5′-^{32}P-radiolabeled, with its 2′,3′-cyclic phosphate provided by intermolecular cleavage of the 5′-labeled RNA with a 10–23 deoxyribozyme (5). For the cleavage assays, the RNA substrate was 5′-^{32}P-radiolabeled.

Selection procedure and cloning

The complete selection procedure (including cloning) is described in detail in our previous report (1). The selection strategy and the activity progression by round are depicted in Figure 2. Following selection round 7, primers analogous to primers 1 and 2 that were used during the selection procedure (but without the 5′-tail on primer 1 or the 5′-phosphate on primer 2) were used to PCR amplify the step C gel band containing the ligation product. The resulting DNA was cloned using an Invitrogen TOPO TA kit as described (1). Note that there is a PCR amplification step immediately prior to cloning. Therefore, if the selection band is composed of subpopulations of molecules that are similar in sequence (e.g. due to convergence on a small number of deoxyribozyme sequences during selection), then it is expected that many clones from that selection will be similar in sequence. As a consequence, similarity of sequences obtained from a single cloning experiment is not evidence for or against

*To whom correspondence should be addressed. Tel: +1 217 244 4489; Fax: +1 217 244 8024; Email: scott@scs.uiuc.edu
Present address:
Amber Flynn-Charlebois, Department of Chemistry and Physics, William Paterson University of New Jersey, Wayne, NJ 07470, USA

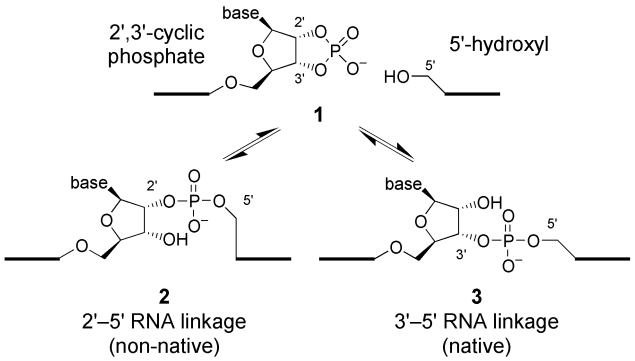

Figure 1. RNA ligation by attack of a 5′-hydroxyl on a 2′,3′-cyclic phosphate. Two isomers are possible: the non-native 2′–5′ linkage (**1**→**2**) or the native 3′–5′ junction (**1**→**3**). For all reported Mg^{2+}-dependent deoxyribozymes that mediate reaction of **1**, the product is **2** and not **3**.

contamination as an origin of the sequence information (see Discussion).

Surveying the kinetic activities of deoxyribozyme clones before sequencing

After cloning, the RNA ligation activities of individual deoxyribozymes were assayed to determine which to sequence and study in more detail. In our initial report (1), this surveying was initiated by PCR amplifying the appropriate region of plasmid miniprep DNA for each clone. This provided enough of each deoxyribozyme to permit ligation assays using the trimolecular format of Figure 4A. However, this method is rather laborious and here we used a related approach that is derived directly from the selection procedure itself. From the plasmid miniprep DNA for each clone, PCR (50 µl with Taq polymerase) was used to amplify the deoxyribozyme strand with incorporation of [α-^{32}P]dCTP, such that the product was internally radiolabeled. This is similar to step C of Figure 2, except that here the template for the PCR amplification is a plasmid DNA. The desired single-stranded product (already bearing a 5′-phosphate from PCR primer 2) was isolated by PAGE and ligated with the right-hand RNA substrate (10 µl reaction volume, 40 pmol right-hand RNA substrate, 50 µM ATP, 10 U T4 RNA ligase, 13 h). This is similar to step A of Figure 2A. Finally, the kinetic activity of the particular deoxyribozyme was assayed at 40 mM Mg^{2+}, pH 9.0, and 37°C, using the RNA–DNA chimeric internally radiolabeled product (~5 pmol) in a 10 µl assay with excess unradiolabeled left-hand RNA substrate (~20 pmol). Time points were taken at 0, 40 and 90 min, with analysis by 8% PAGE; typically 14 clones were assayed in parallel in each experiment. These data provided sufficient information for subjective choice of which deoxyribozymes to study further. We generally pursued those DNA enzymes with at least 40% ligation yield by the 90 min time point.

Kinetics assays

For each deoxyribozyme prepared by solid-phase synthesis, the ligation assays used the trimolecular format shown in Figure 4A. The ^{32}P-radiolabeled left-hand RNA substrate L was the limiting reagent relative to the deoxyribozyme E and right-hand substrate R. The ratio L:E:R was ~1:3:6 to 1:10:30,

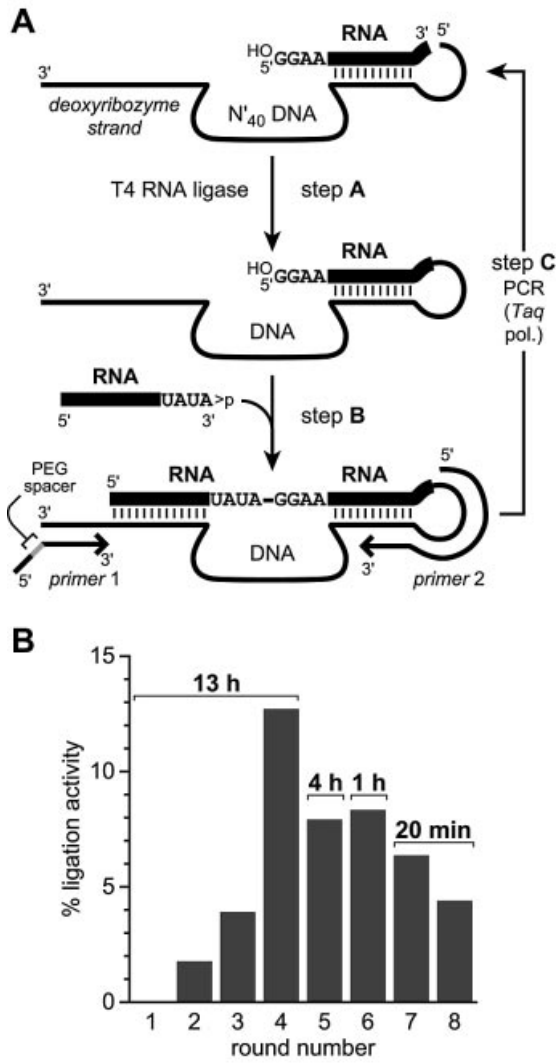

Figure 2. The selection procedure. (**A**) Overview of the strategy. RNA is shown as a thick line and DNA as a thin line. The products are separated by denaturing PAGE after each of the three steps A–C of each selection round. The PEG spacer in PCR primer 1 for step C causes the two single-stranded products to have different lengths, thus permitting their separation by PAGE. A detailed description of the overall selection procedure has been published (1). Here, the reaction conditions during the key step B were 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl and 40 mM $MgCl_2$ at 37°C. (**B**) Progression of ligation activity. For each round, the incubation time for selection step B is indicated; the selection pressure was increased by decreasing the incubation time. The round 7 pool was cloned. Repeating round 5 with a 13 h step B incubation led to 18–21% ligation activity, but we did not further pursue this branch of the selection tree as our principal goal was faster deoxyribozymes.

with the concentration of E equal to ~0.5–3 µM. Increasing the concentration of E or R (or both) did not significantly change the observed k_{obs} values or yields, indicating that these were not limited by availability of E or R (data not shown). See our earlier report (1) for a detailed description of the methods of sample preparation and ligation analysis. Where tested, essentially the same results were observed in the trimolecular format when the limiting reagent was internally radiolabeled R prepared by transcription with [α-^{32}P]CTP (data not shown). In all cases, values of k_{obs} and final yield were obtained by

fitting the yield versus time data directly to first order kinetics; i.e. yield = $Y(1 - e^{-kt})$, where $k = k_{obs}$ and Y = final yield. In all of the curve fits, the data clearly level off at a final yield less than 100%, indicating that the fitted value of Y is meaningful. For the cleavage assays of Figure 6, the ratio of RNA substrate L+R to deoxyribozyme was 1:5.

As tabulated in the Supplementary Material, the left-hand DNA binding arm (i.e. the DNA binding arm that interacts with the left-hand RNA substrate) had up to three mutations relative to the parent sequence. For all assays described here, this DNA binding arm was restored to the fully complementary parent sequence in the deoxyribozymes that were prepared by solid-phase synthesis. For 7Z81 and 7Z48, several of the mutations found in the DNA binding arms of individual clones were demonstrated not to change the ligation rate or yield significantly (data not shown), indicating that these mutations are unimportant. However, for 7Z101, up to a 2-fold increase in k_{obs} was observed when the mutated nucleotide was retained (data not shown). We have not further pursued these relatively modest differences in 7Z101 activity.

RESULTS

Selection strategy

The new selection for improved RNA ligation rate used the general strategy shown in Figure 2A (1). The deoxyribozyme strand was prepared by solid-phase synthesis with a partially randomized N'_{40} region. Each nucleotide N′ of the N'_{40} region had a 25% average chance of incorporating a different nucleotide than is found in the parent 9A2 sequence (1), with an equal probability of including any of the other three standard DNA nucleotides. Each selection round comprised three steps A–C as follows: (A) efficient T4 RNA ligase-catalyzed joining of the right-hand RNA substrate to the deoxyribozyme strand, which contained the randomized DNA region; (B) selection for DNA enzyme sequences that are competent to ligate the left-hand RNA substrate to the right-hand RNA substrate; (C) PCR amplification of the successful DNA enzymes, with retention of the desired single-stranded products for input into step A of the next selection round. Steps A–C were each culminated by denaturing PAGE purification of the desired products. Rounds were performed with increasing selection pressure via reduced incubation time in step B until the activity of the DNA pool reached a plateau, as shown in Figure 2B.

Identification of three new deoxyribozymes and determination of their RNA ligation activities

Individual deoxyribozymes were cloned from the round 7 pool and their ligation activities were surveyed. To permit rapid screening of a large number of DNA enzymes, this survey was initially done using samples obtained by PCR amplification from plasmid miniprep DNA of each clone. Deoxyribozymes with promising RNA ligation activities as judged subjectively by rate and yield were further analyzed by automated sequencing. The resulting sequences segregated largely into three families, with representative members 7Z81, 7Z48 and 7Z101 (Fig. 3; see Supplementary Material for full sequence tabulations).

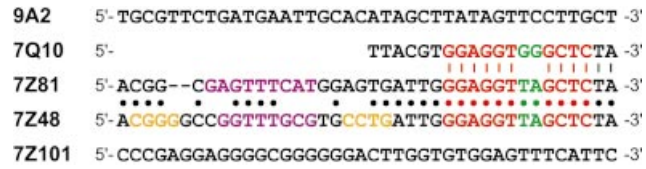

Figure 3. Sequences of the new 7Z81, 7Z48 and 7Z101 deoxyribozymes. The sequences of the 9A2 and 7Q10 deoxyribozymes are provided for comparison (1,2). Identity between 7Q10 and the new DNA enzymes is marked with vertical lines. The red nucleotides constitute the 7Q10-like catalytic core; other colors are used as shown in Figure 5. Identity between 7Z81 and 7Z48 is marked with bullets.

The RNA ligation rate (k_{obs}) and yield of each of these DNA enzymes were characterized in the trimolecular assay format of Figure 4A. Representative data are shown in Figure 4B and the quantitative k_{obs} and yield values from Figure 4C are summarized in Table 1. These assays used deoxyribozymes prepared by solid-phase synthesis, which confirmed both the sequences of the DNA enzymes and the reproducibility of their RNA ligation activities upon independent synthesis.

The ligation assays depicted in Figure 4 were performed under the standard reaction conditions of 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl and 40 mM $MgCl_2$ at 37°C. Because some of the slower DNA enzymes required many hours to achieve maximal ligation under these conditions, we also performed ligation assays in CHES buffer at pH 9.0, where the k_{obs} was higher (Table 1; note that the units on k_{obs} are h^{-1} at pH 7.5 and min^{-1} at pH 9.0). Generally, the ligation yields, as opposed to rates, were comparable at pH 7.5 and pH 9.0 for any particular deoxyribozyme. To expedite the analyses, all of the comparative ligation assays described below were performed at pH 9.0.

The 7Z81 deoxyribozyme has a 7Q10-like core with additional sequence elements

As denoted by red nucleotides in Figure 5, the 7Z81 deoxyribozyme shares the conserved catalytic core stem–loop of the previously reported 7Q10 DNA enzyme (2,3). However, 7Z81 has a considerable number of additional nucleotides in the connecting region between the 7Q10-like catalytic core and the DNA binding arm that interacts with the right-hand RNA substrate. The presence of these extra nucleotides compared with the shorter 7Q10 is as expected from the 40 nt randomized region used in the present selection (Fig. 2A), versus the shorter enzyme region used in the discovery of 7Q10 itself (2). The extra nucleotides found in 7Z81 are not predicted by mfold (6,7) to adopt any specific secondary structure.

The 7Z81 deoxyribozyme has about a 14-fold faster k_{obs} and a higher ligation yield than does 7Q10 (Table 1 and Fig. 4C). The extra nucleotides in 7Z81 are likely responsible for its improved ligation activity. Therefore, a series of deletion mutants, a term that we contract to 'deletants', was used to determine which of these nucleotides are functionally important. Five deletants, 7Z81d1–7Z81d5, were prepared, each missing four specific nucleotides as indicated in Figure 5B. Deletants 7Z81d1, 7Z81d4 and 7Z84d5 are significantly less active than the parent 7Z81 (Table 2), suggesting an important functional role for the corresponding nucleotides. However,

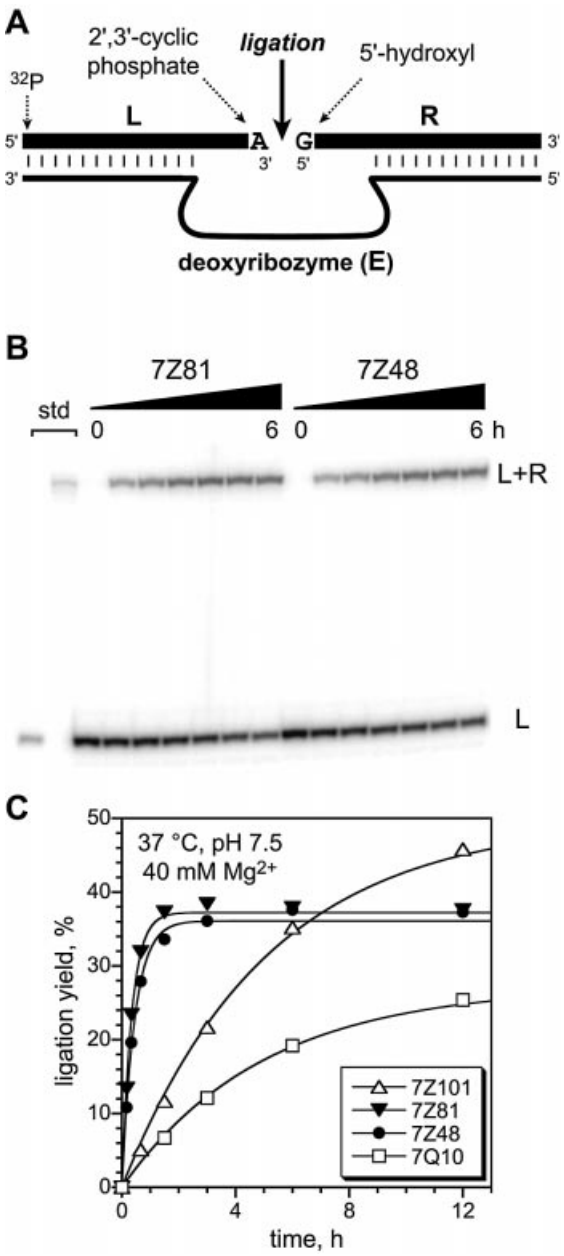

Figure 4. Representative RNA ligation assays for the new deoxyribozymes. (**A**) Trimolecular format of the ligation assays. (**B** and **C**) Ligation assays for the new DNA enzymes under standard incubation conditions (50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, 37°C). Values of k_{obs} from the curve fits are collected in Table 1, along with similar data obtained in CHES, pH 9.0. Data points for 7Z101 and 7Q10 recorded at 24 h fall on the curve fit lines.

the deletants 7Z81d2 and 7Z81d3 are only moderately perturbed in activity relative to the full-length 7Z81. Surprisingly, the 8 nt deletant 7Z81d(2+3) is as active as the parent sequence.

To complete the analysis of 7Z81, we created several additional deletants to more precisely identify the dispensable nucleotides. We found that one additional G nucleotide may be removed from 7Z81d(2+3) without lowering the ligation activity; further deletion leads to a significant drop in ligation yield. The final optimal variant is simply termed 7Z81d (Fig. 5B). We conclude that 7Z81d requires 16 nt in the connecting region between the 7Q10-like catalytic core stem–loop and the right-hand binding arm. Curiously, 8 of these 16 nt are guanosines; i.e. this apparently unstructured connecting region is highly G-rich.

The 7Z48 deoxyribozyme also has a 7Q10-like core

As for 7Z81, comparison of the predicted 7Z48 secondary structure with that of 7Q10 reveals significant similarity (Fig. 5). The 7Z81 and 7Z48 deoxyribozymes also have some common nucleotides outside the 7Q10-like core (Fig. 3). However, 7Z48 has an additional mfold predicted stem–loop within the connecting region (colored yellow in Fig. 5C), in contrast to 7Z81, which has no specific secondary structure predicted for these extra nucleotides. Despite the sequence differences, the ligation rate and yield of 7Z48 are comparable to those of 7Z81 (Table 1).

To identify the nucleotides required for 7Z48 activity, we created nine block deletants of the parent full-length sequence. Each deletant was missing 8 of the 40 enzyme region nucleotides in an overlapping sliding window (Fig. 5C). The first deletant was missing nucleotides 1–8; the second was missing nucleotides 5–12; and so on through nucleotides 33–40; these nine deletants are termed 7Z48d1–7Z48d9. In ligation assays, these deoxyribozymes separated into three kinetic classes (Table 2): those with ligation activity equivalent to that of the full-length 7Z48 sequence (d3 only); those with substantially reduced but still detectable ligation activity (d2, d4 and d5); those with essentially no detectable ligation activity (d6–d9). The eight dispensable nucleotides of 7Z48d3 are marked in purple in Figure 5C and lie in the remote part of the second predicted stem–loop extension; this supports the illustrated secondary structure. In further experimental analysis of 7Z48 we began with the shortened 7Z48d3 sequence, which we more compactly term 7Z48d. Similar to 7Z81d, the connecting region nucleotides of 7Z48d between the 7Q10-like catalytic core stem–loop and the binding arm are relatively G-rich (8 of 19 nt).

Comparison of the 7Q10-like catalytic core between 7Q10 itself and 7Z48d reveals a notable difference: 7Z48d has a TA dinucleotide near the catalytic core at the pentaloop positions corresponding to the GG within 7Q10 (compare green nucleotides in Fig. 5). We examined various mutations of these nucleotides, which were previously shown to be important in a study of 7Q10 itself (3). Mutating TA→GG in 7Z48d led to nearly complete disruption of ligation activity (Table 2). In contrast, the corresponding GG→TA mutations in 7Q10 permitted significant activity, but the yield was lowered slightly compared with the parent sequence (k_{obs} = 0.081 min^{-1} and 21% yield at pH 9.0; data not shown). These data support a functional interplay between the pentaloop nucleotides and the remainder of the deoxyribozyme structure, as separately concluded in our earlier study of 7Q10 (3).

The second putative stem–loop extension of 7Z48d has loop sequence GCCTG (Fig. 5C; note that the purple nucleotides are deleted in 7Z48d). We examined the two 7Z48d mutants GTTCG and ACCTA, where the underlined loop nucleotides were changed via transitions (T↔C or A↔G). The results in Table 2 indicate that the central 3 nt of this loop are relatively unimportant for ligation activity, whereas the flanking 2 nt are required.

Table 1. Summary of ligation and cleavage rates and yields for the new DNA enzymes[a]

Deoxyribozyme	Ligation at pH 9.0 k_{obs} (min^{-1})	Yield (%)	Ligation at pH 7.5 k_{obs} (h^{-1})	Yield (%)	Cleavage at pH 7.5[b] k_{obs} (h^{-1})	Yield (%)
7Q10	0.042 ± 0.002	25	0.19 ± 0.02	27	0.10 ± 0.02	[c]
7Z81[d]	0.60 ± 0.04	35	2.65 ± 0.21	38	0.42 ± 0.02	56
7Z48[e]	0.52 ± 0.01	37	2.13 ± 0.04	37	0.56 ± 0.05	48
7Z101	0.033 ± 0.001	45	0.20 ± 0.01	50	0.014 ± 0.003	[c]

[a]See Materials & Methods for details of the ligation and cleavage procedures. Incubation conditions were 50 mM HEPES, pH 7.5, or CHES, pH 9.0, as indicated, in addition to 150 mM NaCl, 2 mM KCl and 40 mM $MgCl_2$ at 37°C. See Figures 5 and 10 for representative data. The parent 9A2 sequence has ligation k_{obs} = 0.023 min^{-1} at pH 9.0 with 55% yield, but none of the new DNA enzymes has significant sequence similarity to 9A2 (Fig. 3). All values of k_{obs} are averages ± SD of between two and four independent experiments. Yields are generally reproducible to ±5% and are the extrapolated final yields at completion of the reactions.
[b]Data for cleavage of 2′–5′ linked RNA (see text).
[c]The cleavage yield could not be determined accurately due to the low value of k_{obs}. At 12 h the RNA was 24% cleaved (7Q10) or 19% cleaved (7Z101).
[d]The data are for the fully active 7Z81d deletant of Figure 5B.
[e]The data are for the fully active 7Z48d deletant of Figure 5C.

Finally, we examined more closely the predicted 4 bp stem–loop extension in the connecting region of 7Z48d (bp1–bp4; yellow nucleotides in Fig. 5C). The sole G–T wobble pair at bp2 could be changed to the canonical G–C base pair with no significant change in activity (Table 2). Systematic compensatory mutations were then tested at each of bp1–bp4, using either double transversions (e.g. G–C→A–T) or base exchanges (e.g. G–C→C–G). At bp2, G–C→A–T resulted in a 4-fold decrease in k_{obs} along with a drop in yield (Table 2). For all seven of the other tested nucleotide combinations, the activity was even lower, with <6% ligation yield (data not shown). Therefore, the specific identities of the nucleotides within this putative stem of 7Z48d are functionally important. Even compensatory mutations that retain the potential for Watson–Crick base pairing are not tolerated.

The 7Z101 deoxyribozyme is unrelated to previously known deoxyribozymes

Judging from its sequence (Fig. 3), 7Z101 appears to be unrelated to 9A2, 7Q10 or any of our other previously identified deoxyribozymes. It has several related mfold predicted secondary structures (see Supplementary Material); the relative and absolute energies of these predicted structures vary as a function of the Mg^{2+} concentration input into the mfold algorithm. A substantial number of mutagenesis and deletion experiments would be required to investigate this further. The ligation yield of 7Z101 is ~50%, but it is not particularly fast compared with the other DNA enzymes (Table 1). We have not yet pursued additional experimental characterization of 7Z101.

RNA cleavage (rather than ligation) by the new deoxyribozymes

The three new deoxyribozymes 7Z81, 7Z48, and 7Z101 were assayed for their ability to cleave instead of ligate RNA. All of the new deoxyribozymes are capable of cleaving 2′–5′ linked RNA to some extent (Fig. 6), as is 7Q10 (2). The cleavage rates and yields are summarized in Table 1 alongside the ligation data. The 7Z101 cleavage rate and yield are lower than those of 7Q10, consistent with the observation that 7Z101 is a better RNA ligase than is 7Q10. Somewhat surprisingly, 7Z81 and 7Z48 cleave 2′–5′ linked RNA with yields comparable to those for ligation (or even slightly better) under similar incubation conditions. However, their cleavage rates are still ~4-fold lower than their ligation rates. In contrast to the observed 2′–5′ cleavage activity, none of the three new deoxyribozymes is able to cleave 3′–5′ linked RNA (<5% cleavage; data not shown). This is similar to the inability of 7Q10 to cleave 3′–5′ linked RNA (2).

DISCUSSION

Correlation of primary sequence (structure) with improved ligation rate (function) for the 7Q10-like deoxyribozymes

We have identified three new deoxyribozymes with improved RNA ligation activity relative to the previously reported 7Q10 DNA enzyme (Fig. 4). Two of the new deoxyribozymes, 7Z81 and 7Z48, are related to 7Q10 by the presence of a common secondary structural element (Fig. 5). In particular, 7Z81 and 7Z48 each have the core 7Q10-like stem–loop motif (red nucleotides in Fig. 5), but they retain all ~40 random region nucleotides between their DNA binding arms (Fig. 3). In addition, 7Z81 and 7Z48 are over an order of magnitude faster at RNA ligation than is 7Q10 itself (Table 1).

It is clear that the 7Q10-like core is strictly required for ligation activity, because deletions within this core are not tolerated at all (e.g. 7Z48d6–7Z48d9; Table 2). Beyond the requirement for the 7Q10-like core, the other sequence requirements for 7Z81 and 7Z48 are more difficult to interpret using only the available mutagenesis and deletion data. The extra nucleotides in the optimal 7Z81d variant (Fig. 5B) do not adopt any obvious set of secondary structure interactions. Although the extra nucleotides in the optimal 7Z48d variant are consistent with a 4 bp stem–loop structure (yellow in Fig. 5C), Watson–Crick co-variation experiments generally indicate that the specific nucleotides found in the parent sequence are required. Therefore, the conclusion of a Watson–Crick stem–loop structure is only tentative. A more detailed understanding of structure–function relationships in these deoxyribozymes, including determination of the importance of the G-rich connecting regions in 7Z81 and 7Z48, would be aided by direct structural characterization via X-ray crystallography or NMR spectroscopy. Such structural studies are also expected to illuminate the numerous hints of functional

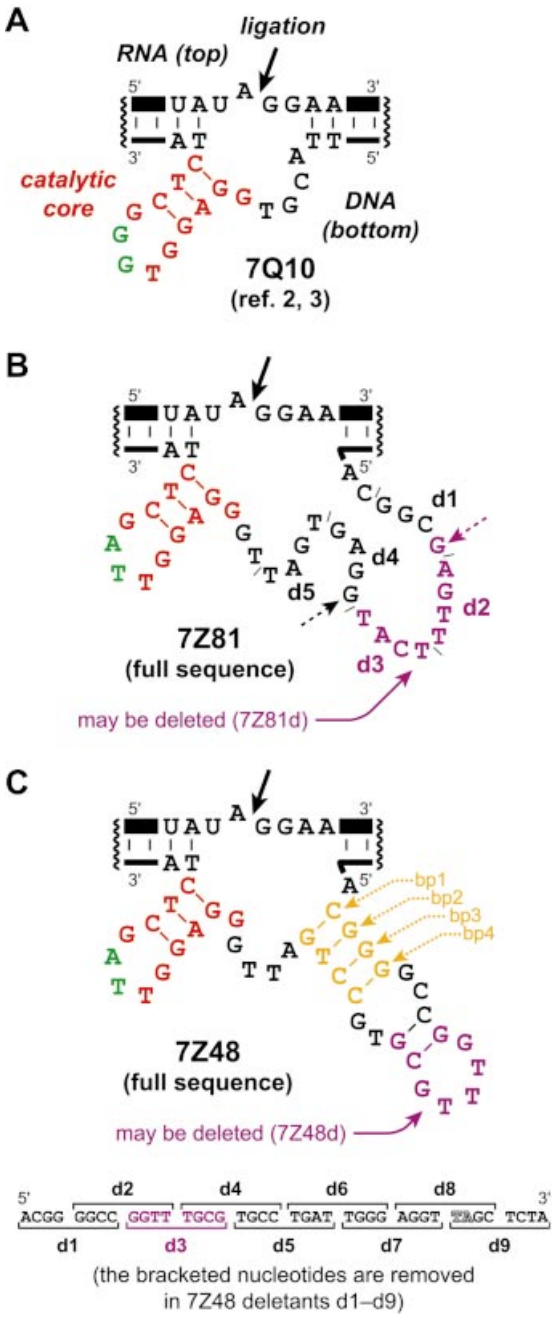

Figure 5. Secondary structure comparison of the previously reported 7Q10 DNA enzyme and the new 7Z81 and 7Z48 DNA enzymes. (**A**) The 7Q10 DNA enzyme (2,3), with the catalytic core nucleotides in red. (**B**) The 7Z81 DNA enzyme. The indicated d1–d5 regions on 7Z81 denote the 4 nt deletion mutants (deletants) assayed for ligation activity. The purple nucleotides in the large right-hand loop may together be removed without disrupting activity [the 7Z81d(2+3) deletant in Table 2]. The G nucleotides marked with dashed arrows were additionally tested by deletion in the context of 7Z81d(2+3). As marked, only one of these two G nucleotides may be deleted while retaining full activity. Additional deletions are also not tolerated (Table 2). (**C**) The 7Z48 DNA enzyme. The purple nucleotides in the large right-hand loop may be deleted without disrupting ligation activity (7Z48d3 deletant in Table 2), which is consistent with the illustrated secondary structure. The parent full-length 7Z48 sequence is shown, along with deletants d1–d9. All secondary structures were predicted by mfold (6,7), which is available online at www.bioinfo.rpi.edu/applications/mfold and also at www.idtdna.com.

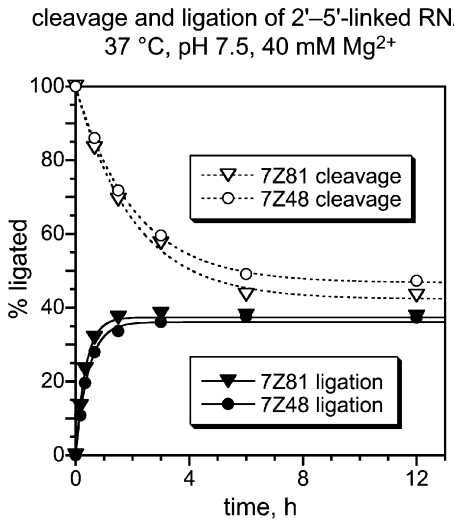

Figure 6. Cleavage of 2′–5′ linked RNA by the new deoxyribozymes, shown with ligation assay data from Figure 4C for comparison. The cleavage assays were performed under the standard incubation conditions (50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, 40 mM $MgCl_2$, 37°C) with a 1:5 ratio of (L+R):E. Values of k_{obs} are collected in Table 1. In parallel assays, a 3′–5′ linked RNA was not cleaved detectably by any of the deoxyribozymes (<5%; data not shown).

interplay between specific structural elements of the deoxyribozymes, as described here and in Ricca *et al*. (3).

The cleavage rate of the hammerhead ribozyme depends strongly upon the presence of tertiary contacts mediated by stem–loop structures that are distinct from its catalytic core (8). We hypothesize that a related consideration applies to 7Q10 as compared with the new 7Z81 and 7Z48 deoxyribozymes, because the latter DNA enzymes also have additional sequence elements that are implicated in improved catalytic function. However, there are no naturally occurring examples of co-varying non-catalytic structural elements for 7Z81 and 7Z48, as exist for the hammerhead ribozyme (8), and only limited data are available from the 7Z81 and 7Z48 artificial phylogenies (see Supplementary Material). Therefore, this hypothesis is best investigated through direct structural analysis of 7Z81 and 7Z48, which is beyond the scope of this report.

RNA cleavage versus ligation: more than merely equilibration?

The RNA ligation reaction involving a 2′,3′-cyclic phosphate and 5′-hydroxyl is reversible (**1**⇌**2**; Fig. 1). According to the principle of microscopic reversibility, any deoxyribozyme that ligates these two RNA substrates (**1**→**2**) is capable of cleaving the product (**2**→**1**), although this cleavage does not necessarily occur with a high rate constant. The new DNA enzymes are indeed capable of RNA cleavage (Fig. 6). To address the mechanism of RNA ligation, it is worthwhile to consider if the new DNA enzymes merely shift the position of equilibrium between **1** and **2** or if the observations require a more complex mechanistic explanation. For 7Z81 and 7Z48, the cleavage and ligation yields are both significant (~50–55 and 35–40%, respectively; Fig. 6 and Table 1), with cleavage favored slightly over ligation. The direct comparison of the cleavage

Table 2. Ligation rates and yields for deletants of the 7Z81 and 7Z48 DNA enzymes[a]

DNA enzyme	Ligation rate (k_{obs}, min^{-1})	Ligation yield (%)
7Z81 deletants		
7Z81 full sequence	0.50	34
7Z81d1	nd[b]	<5
7Z81d2	0.28	26
7Z81d3	0.17	21
7Z81d4	nd	<5
7Z81d5	nd	<5
7Z81d(2+3)	0.60	38
7Z81d(2+3)–G (=7Z81d)	0.60	35
7Z81d(2+3)–2G	0.28	13
7Z81d(2+3)–4nt	nd	<5
7Z81d(2+3)–6nt	nd	<5
7Z48 deletants		
7Z48 full sequence	0.46	38
7Z48d1	nd	[c]
7Z48d2	nd	<5
7Z48d3 (=7Z48d)	0.52	37
7Z48d4	nd	<5
7Z48d5	nd	<5
7Z48d6	nd	[c]
7Z48d7	nd	[c]
7Z48d8	nd	[c]
7Z48d9	nd	[c]
7Z48d GCCTG → GTTCG	0.28	33
7Z48d GCCTG → ACCTA	nd	<5
7Z48d TA → GG[d]	nd	[c]
bp2 G–C[e]	0.49	36
bp2 G–C → A–T	0.12	25
bp2 G–C → C–G	nd	[c]

[a]The incubation conditions were 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl and 40 mM $MgCl_2$ at 37°C.
[b]nd, not determined due to the low yield.
[c]Ligation activity was not detected (<1%).
[d]The indicated green nucleotides of Figure 5C were mutated.
[e]Refers to bp2 as shown in yellow in Figure 5C. 7Z48d itself has bp2 = G–T.

and ligation progress curves as shown in Figure 6 indicates that both types of assay proceed to a common end-point, suggesting that equilibrium is reached from either direction. If this were simply an equilibration with K_{eq} = [ligated]/[cleaved] ≈ 1 (or a little less) as suggested by the end-point, then one would expect the ratio k_{obs}(ligation)/k_{obs}(cleavage) to be ~1 as well (9). However, for both 7Z81 and 7Z48, the ratio k_{obs}(ligation)/k_{obs}(cleavage) is ~4–6 (Table 1). Therefore, some mechanistic explanation beyond mere equilibration must apply, as we independently concluded for 7Q10 on similar grounds (2). The specific nature of this explanation is unclear based on the available data.

Origin of the new deoxyribozyme sequences

In the re-selection starting with the 9A2 sequence we used 25% randomization per nucleotide position, which is typical for re-selections (10). Of DNA molecules in the initial pool, a substantial subpopulation has significantly more randomization than average (see Supplementary Material) and it is possible that such a subpopulation may come to dominate the final enzyme sequences, particularly when strong selection pressure is applied (2). On statistical grounds, one may question the likelihood of obtaining 7Q10-like sequences unrelated to 9A2 (such as 7Z81 and 7Z48) from a re-selection procedure starting with a partially randomized 9A2 sequence. The 7Z81 and 7Z48 sequences both differ from 9A2 at 30 out of 40 nucleotide positions, which is no greater than expected by chance for independent sequences. The probability of having 30 differences out of 40 positions with 25% randomization is only ~4×10^{-11}. However, with 200 pmol ≈10^{14} molecules in the initial selection pool, this still provides ~5000 molecules that differ from the parent 9A2 sequence at 30 or more positions (see Supplementary Material for detailed calculations). If it is allowed that a handful, e.g. 5, of these 30 mutations may have been generated by Taq polymerase during the selection process, rather than being present at the start of selection, then only a smaller number, e.g. 25, of the mutations relative to the parent sequence must have been present within molecules in the initial pool. The probability of 25 or more out of 40 nucleotide differences with 25% randomization is 6×10^{-7}. Therefore, ~70 000 000 molecules in the 200 pmol starting pool had this many mutations relative to the parent sequence. On the basis of such statistical considerations, we cannot exclude the possibility that 7Z81 and 7Z48 emerged directly from the 9A2 re-selection.

On the other hand, it is conceivable that the shorter 7Q10 parent deoxyribozyme or a variant from the laboratory environment contaminated our selection pools and subsequently (via some unspecified mechanism) acquired extra nucleotides to make ~40, as demanded by the size-based selection procedure. However, it is implausible that 7Z101 arose via such contamination, regardless of the origin of 7Z81 and 7Z48, because 7Z101 is unrelated to 7Q10, 9A2 or any other deoxyribozyme sequenced in our laboratory. Ultimately it is impossible to trace the detailed origin of the emergent clones without committing an inordinate amount of time and effort. In contrast, both the sequence and functional data for 7Z81 and 7Z48 are unambiguous. The RNA ligation activities of these two new DNA enzymes surpass those of 7Q10 (Table 1) and there are clear sequence and structural relationships (Fig. 5).

CONCLUSION

Through *in vitro* selection we have identified two new DNA enzymes, 7Z81 and 7Z48, that have the catalytic core motif of the previously reported 7Q10 deoxyribozyme. The additional nucleotides in the enzyme regions of 7Z81 and 7Z48 relative to 7Q10 are responsible for their order of magnitude increase in RNA ligation rate. Thus a structure–function relationship with regard to RNA ligase activity is found in the collection of 7Q10, 7Z81 and 7Z48 deoxyribozymes. These data form the basis for future experiments aimed at a more detailed structural understanding of DNA-mediated RNA ligation reactions. Additionally, the results directly demonstrate how larger enzyme regions enable superior functional activity for DNA enzymes (11).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

This research was supported by the Burroughs Wellcome Fund (New Investigator Award in the Basic Pharmacological Sciences), the March of Dimes Birth Defects Foundation (Research Grant no. 5-FY02-271), the National Institutes of Health (GM-65966), the American Chemical Society Petroleum Research Fund (38803-G4) and the UIUC Department of Chemistry (all to S.K.S.). S.K.S. is the recipient of a fellowship from The David and Lucile Packard Foundation.

REFERENCES

1. Flynn-Charlebois,A., Wang,Y., Prior,T.K., Rashid,I., Hoadley,K.A., Coppins,R.L., Wolf,A.C. and Silverman,S.K. (2003) Deoxyribozymes with 2′-5′ RNA ligase activity. *J. Am. Chem. Soc.*, **125**, 2444–2454.
2. Flynn-Charlebois,A., Prior,T.K., Hoadley,K.A. and Silverman,S.K. (2003) *In vitro* evolution of an RNA-cleaving DNA enzyme into an RNA ligase switches the selectivity from 3′-5′ to 2′-5′. *J. Am. Chem. Soc.*, **125**, 5346–5350.
3. Ricca,B.L., Wolf,A.C. and Silverman,S.K. (2003) Optimization and generality of a small deoxyribozyme that ligates RNA. *J. Mol. Biol.*, **330**, 1015–1025.
4. Wang,Y. and Silverman,S.K. (2003) Deoxyribozymes that synthesize branched and lariat RNA. *J. Am. Chem. Soc.*, **125**, 6880–6881.
5. Santoro,S.W. and Joyce,G.F. (1997) A general purpose RNA-cleaving DNA enzyme. *Proc. Natl Acad. Sci. USA*, **94**, 4262–4266.
6. SantaLucia,J.,Jr (1998) A unified view of polymer, dumbbell and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci. USA*, **95**, 1460–1465.
7. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
8. Khvorova,A., Lescoute,A., Westhof,E. and Jayasena,S.D. (2003) Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. *Nature Struct. Biol.*, **10**, 708–712.
9. Stage-Zimmermann,T.K. and Uhlenbeck,O.C. (2001) A covalent crosslink converts the hammerhead ribozyme from a ribonuclease to an RNA ligase. *Nature Struct. Biol.*, **8**, 863–867.
10. Knight,R. and Yarus,M. (2003) Analyzing partially randomized nucleic acid pools: straight dope on doping. *Nucleic Acids Res.*, **31**, e30.
11. Lozupone,C., Changayil,S., Majerfeld,I. and Yarus,M. (2003) Selection of the simplest RNA that binds isoleucine. *RNA*, **9**, 1315–1322.